The Bayesian Conditional Independence Model for Measurement Error: Applications in Ecology

R. J. Denham, M. G. Falk and K. L. Mengersen

Abstract

The measurement error model is a well established statistical method for regression problems in medical sciences, although rarely used in ecological studies. While the situations in which it is appropriate may be less common in ecology, there are instances in which there may be benefits in its use for prediction and estimation of parameters of interest. We have chosen to explore this topic using a conditional independence model in a Bayesian framework using a Gibbs sampler, as this gives a great deal of flexibility, allowing us to analyse a number of different models without losing generality. Using simulations and two examples, we show how the conditional independence model can be used in ecology, and when it is appropriate.

1 Introduction

Measurement error (ME) models or errors–in–variables models refer to regression models in which both the predictors and the response are measured with error. The recognition of such models can be traced to Adcock (1877; 1878), who formulated orthogonal least squares to deal with the situation in which both the predictors and the response are measured with error. Thorough reviews of ME models have been compiled by, for example Cheng and Ness (1999) and Fuller (1987). Their use in fields such as agriculture is well documented, and more recently they have been applied in epidemiology. ME models under a Bayesian framework have been proposed by a number of researchers (Clayton 1992; Richardson and Gilks 1993b; Dellaportas and Stephens 1995; Mallick and Gelfand 1996; Müller and Roeder 1997)

However, as Cheng and Ness (1999) note, ME models are more commonly appropriate than is recognised by many analysts. This is certainly true when considering applications to ecological data. In many situations in ecology, the predictor variables, or covariates, are measured with an acknowledged degree of error, yet this is typically not accommodated in traditional regression or generalized linear modelling frameworks. Zidek et al. (1996) note that when a casual variable is measured with error it may be missed with significance shifted to a proxy collinear variable. The conditional independence model is one approach to ME that can accommodate this type of data in certain ecological analyses.

As in Richardson and Gilks (1993b) we implement this in a Bayesian framework. Such an approach "follows exactly the structure of the error problem without recourse to artificial assumptions ..., propogates the resulting uncertainty through to the parameter estimates, ... [and is] flexible" (Richardson and Gilks 1993b, p. 1716). Statistical methods that adjust for covariate measurement error are rarely used in ecology, but see Fernandes and Leblanc (2005) and Yuan (2007) for examples.

ME models in ecology have slightly different goals to those in epidemiology. Often in epidemiology there is one parameter of interest that measures an outcome as a response of exposure or treatment. The goal is for accurate and precise parameter estimation. However, in an ecological context we have the additional goal of prediction, that is, how well can we accurately and precisely predict a new value. In this paper we will provide a fully simulated example, along with two examples in ecology that focus on both estimation and prediction, and models will be compared on the basis of these goals. The first example aims to predict the distribution of a tree species whereas the second aims for accurate parameter estimation so as to explain some true underlying relationship between age and length of a particular fish species.

This paper is structured such that Section 2 introduces Bayesian Conditional Independence Modelling and Section 3 applies the ME model to the three case studies listed earlier. This is followed by a discussion in Section 4 to conclude the paper.

2 Bayesian Conditional Independence Modelling

It is the conditional independence modelling of Richardson and Gilks (1993b) that we will be concerned with in this paper. Their work illustrated how ME models could be applied to accurate and precise parameter estimation in epidemiological studies. A ME model might consist of a **disease status** (response) which is related to **risk factors** (explanatory variables) for each individual, consisting of truly known risk factors C and unknown factors X which are understood only through one or more surrogate measures Z. Here, X represents an $n \times p$ matrix, so that for each component of X, say x_j with $j \in 1 \dots p$ has m surrogates z_{jk} with $k \in 1 \dots m$. Similarly, C is a $n \times l$ matrix of additional risk factors.

Richardson and Gilks (1993b) divided their model into three submodels describing their role in epidemiological situations:

1. *a disease model* expressing the relationship between the risk factors C and X and the disease status y;

2. *a measurement model*, which expresses the relationship between the surrogate measures Z and the true unknown risk factor X;

3. *an exposure model*, which specifies the distribution of the unknown risk factor X in the general population.

This modelling framework can easily be transferred to ecological situations. For example, say we were interested in explaining the relationship between a species and its environment, and predicting a species' distribution, these submodels could be redefined as:

1. *a species' distribution model* expressing the relationship between the explanatory variables X and C and the response. The response would typically represent the presence or absence of the species;

2. *a measurement model*, which expresses the relationship between the surrogate measures Z and the true unknown habitat variable X;

3. *a habitat model*, expressing the distribution of the habitat variables in the environment.

Of course there could be many interpretations of such submodels. So for generality, these submodels will be termed *a regression model*, *a measurement model* and *a prior model*, which can be written as:

regression model	$p(y\|X, C, \beta)$
measurement model	$p(Z\|X, \lambda)$
prior model	$p(X\|\pi)$

where C, X, y, Z are as previously defined and β, λ, π are model parameters. It is important to note that in this formulation y and Z are independent conditional on X. In other words, if X is known, the addition of Z adds nothing to our knowledge of y.

In order to formulate the relationship between the explanatory variables X and the surrogates Z, some data linking X and Z is required. This is often in the form of a *validation set* in which a number of records have both X and Z measured. Because measurement of the true value X is usually difficult or costly, the validation set is usually small in comparison to the rest of the data. The validation data is often categorised into *internal* or *external* data. For the purposes of this paper, we define an internal validation data set as one in which the response y is captured along with X and Z. The alternative external validation set then has information only on X and Z. This definition follows that of Richardson and Gilks (1993b). Other researchers define the differences between external and internal slightly differently. Carroll et al.. (1995), for example, define an internal validation data set as a subset of the data collected for the main study, while an external validation set would be the data collected from a separate sampling exercise. Kuha (1997) formalise the definition further by dividing the data sets into a primary and a validation set, with an internal validation set defined as one in which the prior model (or exposure model) is the same as in the primary data. Kuha (1997) also emphasises the requirement that the measurement model be the same in both the validation and primary data.

Alternatively, the relationship may be established by *repeated determinations* of X by one or more surrogate methods. The assumption here is that while we

cannot directly measure X, multiple measurements of an unbiased estimator will provide information on X. An application could include a combination of validation and repeated measures.

To keep notation simple, we describe the model using a single known risk factor c, a single unknown risk factor x, with a single surrogate z. The joint distribution can be written as:

$$p(\beta)p(\lambda)p(\pi)\prod_i^n p(x_i|\pi)\prod_i^n p(z_i|x_i,\lambda)\prod_i^n p(y_i|x_i,c_i,\beta).$$

The model can be appreciated graphically using a Directed Acyclic Graph (DAG) (Figure 1). By convention, square boxes represent known quantities and round nodes represent random variables. Thus the validation set includes the x_i in the square node and the adjacent z_i in the square node. Again note that the link, represented by arrows, between the surrogates z and the response y is only through the parameter x. The direction of the arrows reveal the nature of the relationship between the parameters. Classic error models specify the conditional distribution of surrogates z and true x values. The simplest example of a classic error model involves an additive error, $z = x + u$, where u represents the measurement error. We can interpret this as z being a degraded version of x, shown in the graph by arrows pointing from x to z. An alternative error model, the Berkson model, uses fixed z values, with x values varying. Under a Berkson error model, the arrows between x and z would be reversed (Berkson 1950).

Notice that the graph includes two symbols for x, a square box within the validation set, and a round node below. This illustrates that we are treating the unobserved x_i values as parameters to be estimated. Under a Bayesian framework, these parameters are considered random variables. Thus, the validation data provides information on the relationship between x and z via λ, and this information is used in the main study to strengthen the information on the relationship between y, x and c through β. Richardson and Gilks (1993a) provide further examples and interpretation of graphs for measurement error models.

3 Examination of the ME Model

Here we provide examples of the ME model and evaluate its performance in terms of accuracy and precision in explaining and predicting variables.

3.1 Combining Low and High Quality Data

Consider a situation in which we have access to few high quality data points, but relatively abundant low quality data. Data of this form often prove difficult for users to adequately use, and simplifications are often made. For example, when it is assumed that z is an unbiased degraded version of x, the differences are often ignored and the x and z values combined into a single explanatory variable. Alternatively, quality focused users may reject the z values as being

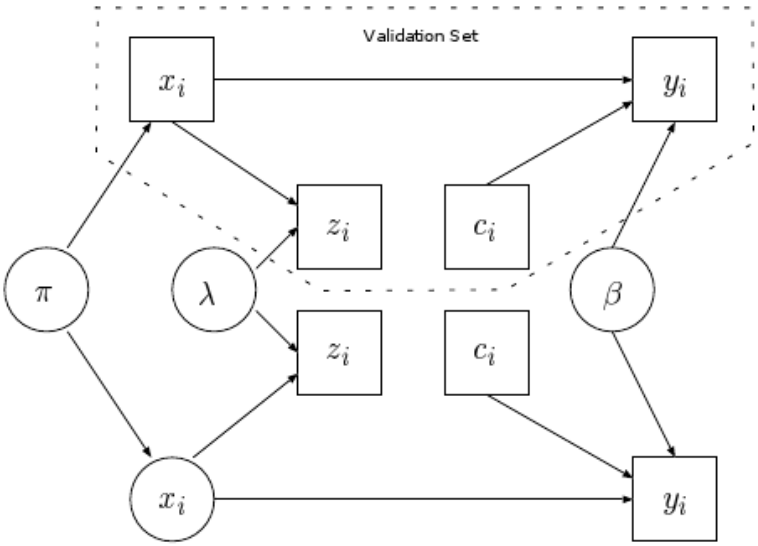

Figure 1: Directed Acyclic Graph (DAG) for the measurement error model (Richardson and Gilks 1993a). The validation set is indicated by the box in broken lines.

too unreliable and use only the accurate x values. In this paper, the first strategy will be called the *naïve analysis* and the second the *validation alone analysis*. Both these strategies ignore important information in the data, so a method that makes use of both x and z would be preferable. A simple simulation allows us to compare the naïve analysis, the validation alone analysis and the ME model.

Consider, for example, a dataset comprising of a response variable y and an explanatory variable x. Only 15 records are available for x, but there are 200 records of a variable z, where z is a biased, degraded version of x. The model is as follows:

$$x \sim N(0,1); \qquad y \sim N(\alpha + \beta x, \sigma_y); \qquad z \sim N(\phi + \psi x, \sigma_z), \tag{1}$$

where $\alpha = 1.0$, $\beta = 1.0$, $\sigma_y = 0.25$, $\phi = 0.1$, $\psi = 1.5$ and $\sigma_z = 0.45$.

Closed form posterior distributions can be found for the naïve analysis and the validation alone analysis, however the posterior for the ME model is not straightforward. A similar problem is outlined by Gustafson (2003) who provides the solution. In more complex cases, Bayesian analysis using Markov Chain Monte Carlo could be undertaken using WinBUGS (Spiegelhalter et al.. 1996). For exposition we do that here. Priors were chosen to add as little information as possible. Thus, the priors for the parameters were:

$$x \quad \sim \quad N(\mu_x, \sigma_x)$$

$$
\begin{aligned}
\alpha &\sim N(\mu_\alpha, \sigma_\alpha) \\
\beta &\sim N(\mu_\beta, \sigma_\beta) \\
\phi &\sim N(\mu_\phi, \sigma_\phi) \\
\psi &\sim N(\mu_\psi, \sigma_\psi) \\
\mu_x, \mu_\alpha, \mu_\beta, \mu_\phi, \mu_\psi &\sim N(0, 100000) \\
\frac{1}{\sigma_x^2}, \frac{1}{\sigma_\alpha^2}, \frac{1}{\sigma_\beta^2}, \frac{1}{\sigma_\phi^2}, \frac{1}{\sigma_\psi^2} &\sim \text{Gamma}(0.1, 0.1).
\end{aligned}
$$

For each model three simultaneous chains were used. For the simpler validation alone analysis and the naïve analysis, a burnin of 2,500 iterations per chain were used, with the posterior distributions summarised by a further 2,500 iterations. This appeared sufficient for conversion, as assessed visually and via the potential scale reduction factor (Gelman and Rubin 1992). Convergence for the ME model proved more problematic, with a longer run length of 75,000 used, the first 37,500 of which were discarded.

Table 1: Results of Gibbs sampling under different strategies. Standard deviations for each parameter are given in parentheses.

Parameter	Analytic Values	Validation Alone	Naïve	ME
α	1.00	1.02 (0.09)	0.96 (0.03)	1.08 (0.06)
β	1.00	1.03 (0.10)	0.59 (0.02)	1.02 (0.06)
σ_y	0.25	0.33 (0.07)	0.38 (0.02)	0.27 (0.03)
ϕ	0.10	-	-	0.21 (0.09)
ψ	1.50	-	-	1.60 (0.09)
σ_z	0.45	-	-	0.44 (0.05)

We get some idea of the accuracy and precision from the results as presented in Table 1. The posterior distributions for each parameter are summarised here using the mean value and standard deviation. We note, for example, that the precision of the parameters for the naïve analysis are greater than the two other models, but that they are also less accurate than the other models. The ME model appears to balance the other models in that it is more precise than the validation alone analysis, and more accurate than the naïve analysis.

Figure 2 (a) provides a plot of the data. Since x and z are similar in magnitude, it is possible to plot both quantities on the y axis, but it should be noted that the horizontal axis corresponds to x. Figure 2 (b) shows three regression lines corresponding to the line resulting from parameter fits for each of the three models, which demonstrates that the parameter estimates of α and β from the ME model and the validation alone analysis are close to the true values. These two models are consequently more *accurate* than the naïve analysis.

Figure 3 shows credible and prediction intervals for each model. The value of including z becomes clear when we look at the credible intervals for the

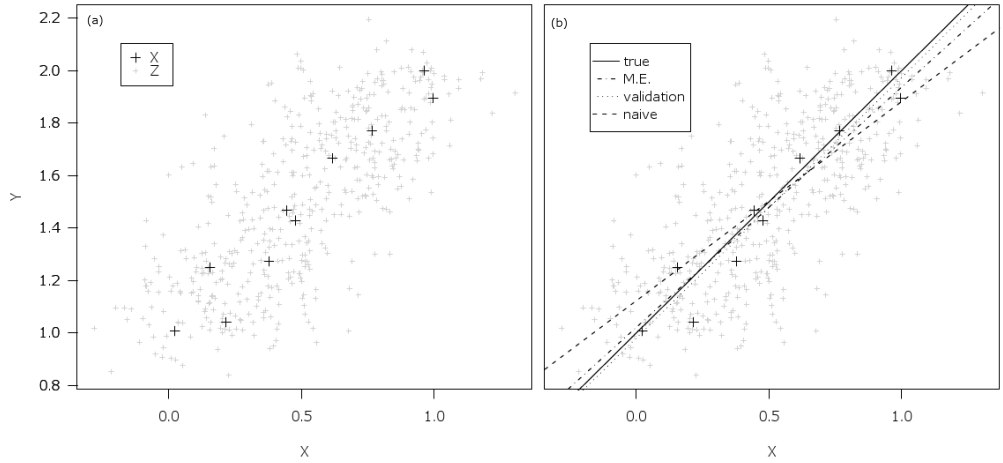

Figure 2: Comparison of regression with mixed data. Panel (a) shows the data, with the accurate (x) values marked by black crosses, and the less accurate (z) values marked by smaller grey crosses. Panel (b) shows the regression lines for the ME model (M.E.), the validation alone analysis (validation), the naïve analysis (naive) as well as the true relationship (true).

expected value of y given a new x, denoted here by $\bar{y}|\tilde{x}$. The extra data allows a more *precise* estimate of the parameters. Figure 3 (a) reflects this with a relatively narrow credible interval for the naïve analysis. When we consider the credible interval for a new y given a new x, or $\tilde{y}|\tilde{x}$, we take into account the uncertainty due to the parameter estimates and the spread of the points around the regression line. Consequently, the ME model does rather better in this situation, since it uses the large amount of z data to reduce the uncertainty in the parameter estimation, and uses the accurate x value to reduce the uncertainty due to scatter around the line. This is reflected in Figure 3 (b).

For the naïve analysis we are treating x and z as the same variable, so it is not possible to predict based on either x or z alone. However, because there are so many more z observations than x observations, this model is approximately equivalent to ignoring the x values altogether. We can use the ME model to predict a new y from a new z as well, but because we have assumed y and z are independent conditional on x, we now allow for uncertainty in the relationship between x and z when predicting y. This is reflected in the relatively wide credible interval for the ME model for $\bar{y}|\tilde{z}$, given in Figure 3 (c). Even allowing for this uncertainty, the credible interval for $\tilde{y}|\tilde{z}$ in the ME model is still comparable to the naïve analysis (see Figure 3 (d)).

This comparison is however from a single draw from the populations used in the simulation. Does this hold in general? To explore this further, the simulation

was expanded slightly to cover a range of values for σ_z, with 20 simulations for each value of σ_z. The naïve analysis is just simple linear regression, hence we know the posterior predictive distribution for $\tilde{y}|\tilde{z}$ is a t-distribution, with $n-2$ degrees of freedom, mean $\hat{\alpha}+\hat{\beta}\tilde{z}$ and scale $s\sqrt{1+(1,\tilde{z})^T(z^Tz)^{-1}(1,\tilde{z})}$ where s^2 is the standard estimate of σ^2. The analytical distribution can thus be compared with the posterior predictive distribution of the ME model from the Gibbs sampler. Figure 4 shows the comparison for a single value of $\tilde{z} = 0.5$. From these results, there is no suggestion that the posterior predictive distribution $\tilde{y}|\tilde{z}$ for the ME model is any more dispersed than that of the naïve analysis. In other words, we do not lose anything using the measurement error model when we wish to predict using future noisy z values. Table 2 summarises the properties of each of the models.

Table 2: Comparison of models for the simulation data. The rankings in each category range from "good" (* * *) to "poor" (*). NA refers to those quantities not available (the validation alone analysis, for example, does not use the z data, and so quantities such as $\tilde{y}|\tilde{z}$ are not available).

		ME model	naïve analysis	validation alone analysis
Accuracy		* * *	*	* * *
Precision	$\tilde{y}\|\tilde{x}$	* * *	NA	* *
	$\bar{y}\|\tilde{x}$	* *	NA	*
	$\tilde{y}\|\tilde{z}$	* *	* *	NA
	$\bar{y}\|\tilde{z}$	*	* * *	NA

In conclusion, depending on whether we are interested in the relationship between x and y at all, and whether future predictions will be based on x or z, the ME model can be used to improve our analyses. The ME model is relatively accurate and more precise compared to considering x alone. However, the precision when predicting $\bar{y}|\tilde{x}$ is lower than for the naïve analysis, but higher when predicting $\tilde{y}|\tilde{x}$. The precision for predicting $\tilde{y}|\tilde{z}$ is approximately the same for both the naïve analysis and the ME model. Accurate, but expensive data, can be augmented with cheaper, lower quality data as it allows greater certainty in the parameter estimates and improves our predictions for both $\tilde{y}$ and $\bar{y}$.

3.2 Species' Distribution Modelling

In our first example demonstrating the use of the ME model for ecological data, we consider the example of species distribution modelling. Recent work has stressed the importance of *modelling the environmental niche* of the species (Austin and Meyers 1996). In this approach, the use of direct causal environmental variables rather than indirect ones is preferred. For example, altitude has no direct bearing on the occurrence or otherwise of a species, but may be considered a surrogate for temperature, a direct contributor to the establishment and survival of a species. It is possible to generate direct correlates of some of the indirect variables. ANUCLIM for example, has been used to generate

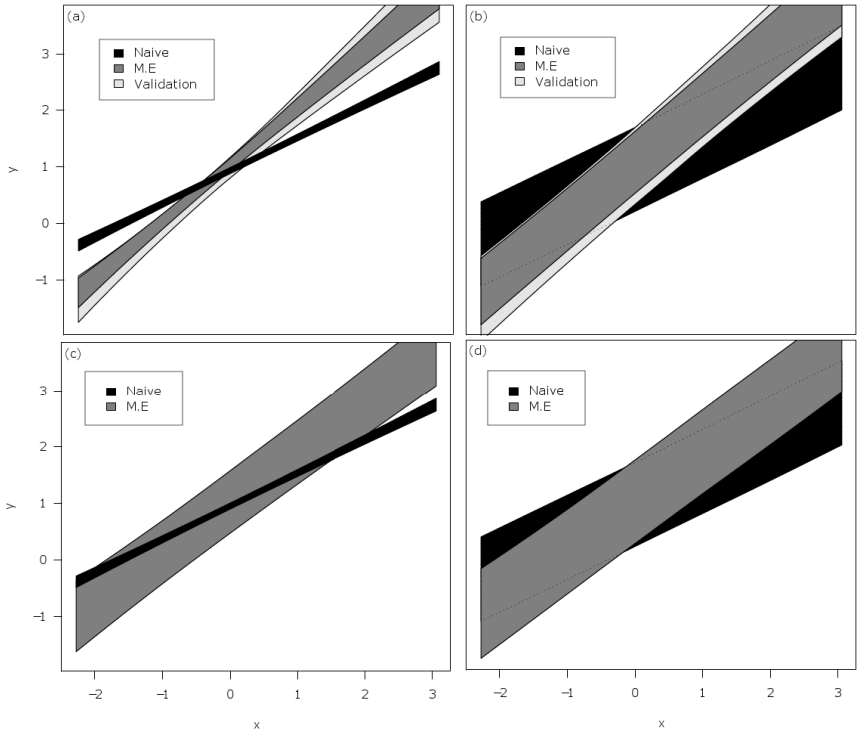

Figure 3: Credible intervals for each model. Panel (a): 95% credible interval for the expected value of y given a new x ($\bar{y}|\tilde{x}$). Panel (b): credible intervals for a new y given a new x ($\tilde{y}|\tilde{x}$). The lower two figures are similar but rather than using a new x, we use a new z. Panel (c): $\bar{y}|\tilde{z}$. Panel (d): $\tilde{y}|\tilde{z}$. The validation alone analysis cannot of course be used with new z values.

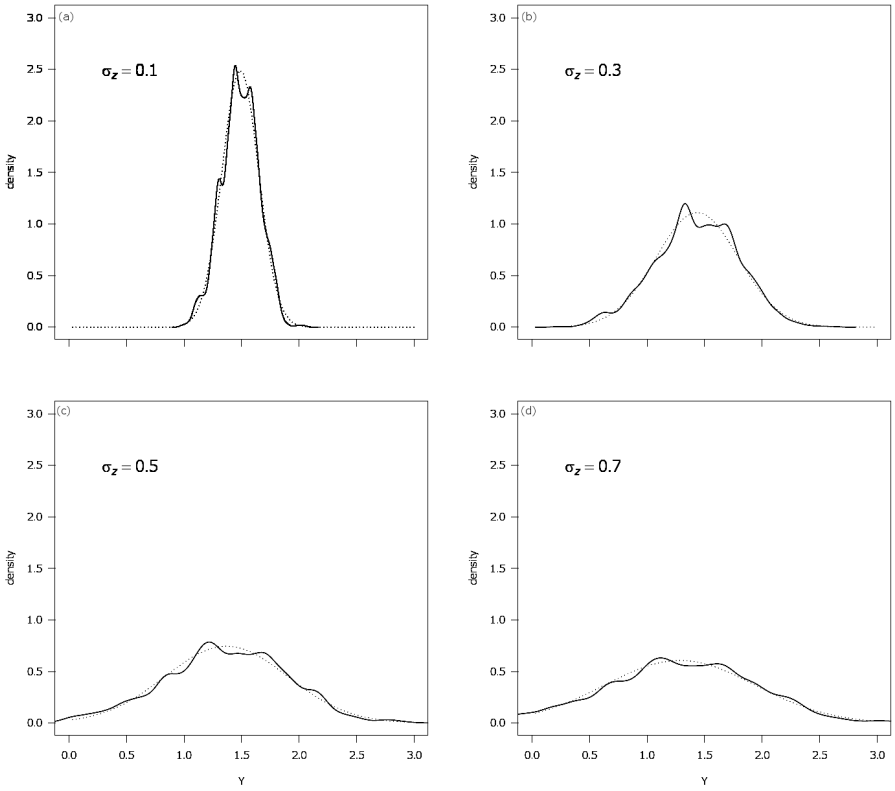

Figure 4: Comparison of $\tilde{y}|\tilde{z} = 0.5$ for the ME model and naïve analysiss. The broken line represents the analytical distribution derived for the naïve analysis, the unbroken line is a summary of the posterior predictive distribution of the ME model. The simulation was repeated 20 times for each of the four values of σ_z.

climatic data given geographic details (Hutchinson et al.. 1998).

The distribution of Gympie messmate (*Eucalyptus cloeziana*), an important timber species in South East Queensland, has been modelled using this approach. The explanatory variables used were extensive, and included 16 climatic indices (rainfall, temperature, light, seasonality), 6 topographic indices (wetness, slope/aspect exposure effect), 4 soil indices (depth, texture, permeability, fertility), and 3 spatial heterogeneity covariates (Williams et al.. 2000).

Of these variables, it is the soil indices which are most difficult to capture, and for the modelling of Gympie messmate were available only as a score based on expert opinion. The score for soil depth indices therefore can be considered a surrogate for true soil depth.

To illustrate how this situation might be modelled using the ME model, we take an extremely simplified dataset, consisting of 1000 observations and just four variables; *annual rainfall*, *mean monthly maximum temperature*, *mean monthly maximum flat surface solar radiation* and *substrate depth index*. The final variable is a surrogate for true substrate depth.

A validation set consisting of a set of 366 observations of which true soil depth and the surrogate were also obtained. The response was not available for the validation set (thus the data consists of an external validation set).

The measurement error model appropriate for this situation is:

$$\begin{aligned} Y_i &\sim \text{Bern}(\alpha_i) \\ \text{logit}(\alpha_i) &= \beta_0 + \beta_1 c_{\text{rain}i} + \beta_2 c_{\text{temp}i} + \beta_3 c_{\text{rad}i} + \beta_4 x_{\text{depth}i} \\ z_i &\sim N(\phi + \psi x_{\text{depth}i}, \theta) \\ x_{\text{depth}i} &\sim N(\mu, \tau). \end{aligned}$$

where Y_i is the presence/absence of Gympie messmate.

Figure 5 provides a graphical representation of this model.

Parameter estimation was via the Gibbs sampler, implemented in the BUGS software (Spiegelhalter et al.. 1996).

Table 3 compares the measurement error model with the standard model. Of note here is the relatively small values for the parameters ϕ and ψ, suggesting a rather poor relationship between x_{depth} and the surrogate z. A plot of the validation set confirms this (Fig 6); the correlation is 0.0075.

The contribution made by the ME model is slight, with a deviance only slightly less than the standard model.

It is really not surprising that the ME model performed poorly in this example, since it is not clear that the assumption of conditional independence is valid in this case. A likelihood ratio test of the ME model and the standard naïve analysis gives a p value of 0.0556 showing some mild evidence that the conditional independence assumption is not empirically verified. Also, the relationship between the surrogate and the true variable is slight, and it is possible that the surrogate adds information not explained by the supposed true underlying parameter. There are a number of reasons why this may occur. Firstly, the quality of the validation set may be in question. In this case, the true

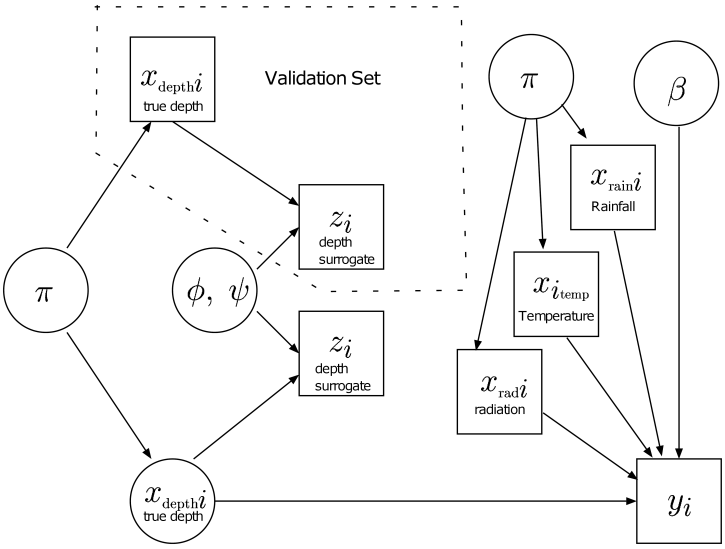

Figure 5: DAG of externally validated ME model.

Table 3: Comparison of the measurement error model with the standard model.

	Standard	ME
β_0	-13.82 (1.65)	-17.77 (5.41)
β_1	-23.73 (2.23)	-29.13 (8.66)
β_2	18.61 (2.08)	25.77 (7.72)
β_3	-29.71 (2.73)	-33.42 (9.42)
β_4	11.59 (1.46)	-2.27 (2.53)
ϕ	–	-0.0068 (0.86)
ψ	–	-0.0069 (0.10)
D	29.91 (11.45)	24.13 (12.14)

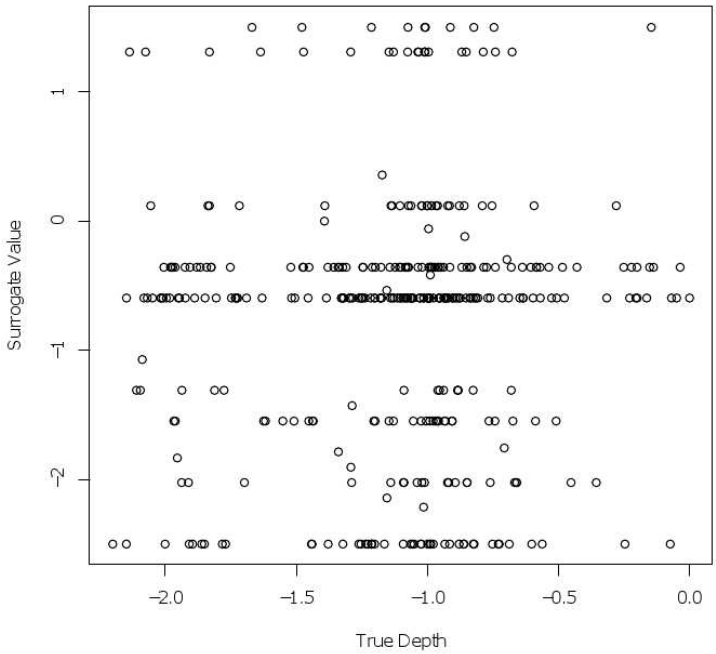

Figure 6: Plot of validation data. There is no clear relationship between the surrogate and the true data. The correlation is 0.0075.

soil depths were collected at point locations, yet the surrogate values were determined from expert opinion at a much coarser resolution. Additionally, the experts provided information on soil suitability for the species in question, and the label *soil depth* was somewhat arbitrary. The true underlying soil property is likely to be more complex than just soil depth, and may include other soil physical properties.

In this situation, the explanatory variable of interest is the less accurate, presumably cheaper data z. The auxiliary data then is the limited amount of accurate, expensive data x. Since the aim is prediction, and in general we wish to predict using the region wide z values, then knowledge of x and its relationship with the response is of little interest. Even the claim that we get closer to the "true" relationship between the response and a set of explanatory variables is spurious here, since the true underlying parameters which determine a species fitness for a particular location are likely to be too complicated to capture. Even the most accurate and detailed measurements are likely to be surrogates for other parameters. In a sense, all these models will be empirical.

3.3 Fish Age and Otolith Measurements

Fish age is a crucial component of fisheries management, needed for calculations of growth rate, mortality rate and productivity. The implications of inaccurate age estimates can be extreme, occasionally contributing to the overexploitation of a population or species. This has been recognised for sometime (Beamish and McFarlane 1983) and Campana (2001) provides a thorough review of the issues. The impact of inadequately accounting for the measurement error of fish age can occasionally have disastrous consequences for some fish populations. Several species have suffered from overexploitation as a result, including the orange roughy (*Hoplostethus atlanticus* Collett) off New Zealand, the *Sebastes* spp fisheries off eastern and western Canada and that of walley pollock (*Theragra chalcogramma* Pallas) in the central Bering Sea. Hence any modelling of fish age must be completed with accurate and precise parameter estimates, to ensure the abundance of fish species in the future.

The most common method of estimating fish age is by counting annual (or daily) growth increment marks on fish otoliths. Unfortunately, there is often some error associated with the conversion from otolith mark counts to fish age. Campana (2001) group the possible sources of error as (a) process error, in which the the laying down of annual increments is not constant throughout the fish's age, or the part of the otolith being examined does not hold a complete record of annual markings; and (b) interpretation error, in which due to the subjective nature of identifying marks on the otoliths there is some variation between readers and laboratories of a given otolith. Assessing the process error can be very difficult without reference data of known age fish, although there are a number of other methods available (see Campana (2001) for details). The interpretation error is often assessed by using multiple readers to estimate ages of a subset of the otoliths available.

To investigate the effects of measurement error, we used simulated data to

model the age-growth curve. This is one of the key management tools for many fisheries, often expressed using the von Bertanlaffy equation:

$$L_t = L_\infty \left\{1 - \exp\left[-k(t - t_0)\right]\right\},$$

where L_∞ is the maximum length that might be attained under ideal conditions, k is the species growth rate and t_0 a negative constant.

This simulation assumes there is no process error, but that there is interpretation error, which still can result in biased parameter estimates.

A common situation is that the interpretation error is larger as the true age increases; see Campana (2001) for further information. To recreate this situation, we designed the simulation as follows:

$$\begin{aligned} x &\sim 1 + \text{Gamma}(\alpha, \beta) \\ y &\sim N(L_\infty \left\{1 - \exp\left[-k(x - t_0)\right]\right\}, \sigma_y^2) \\ z &\sim \text{Gamma}(\lambda, \lambda/x) \end{aligned}$$

where $\alpha = 3$, $\beta = 0.2$, $L_\infty = 130$, $k = 0.15$, $t_0 = -0.05$, $\sigma_y = 7$ and $\lambda = 6$.

The response y is representative of fish length, x values would represent the true (but unknown) ages, and the z values the observed age. The fish lengths were restricted to be greater than zero.

The choice of the gamma distribution to model the interpretation error implies that the error increases as the square of the true age increases, since $E(z|x, \lambda) = x$ and $\text{Var}(z|x, \lambda) = \frac{x^2}{\lambda}$. Figure 7 illustrates the error pattern.

For a subset of the data, we have multiple readings of each of the otoliths. These might be by the same reader or by different readers, but we assume that there is no bias between readers. The mean of the multiple readings would then be an estimate of the true age. For the simulation, we chose the total number of otoliths as $N = 1000$, $n = 200$ had $m = 5$ multiple readings. We used the Bayesian measurement error model to perform parameter estimates, computed in WinBUGS. 100 simulations were run, each with a new random draw for x, y and z. A typical data set is shown in Figure 8.

Suitably vague priors for the parameters were:

$$\begin{aligned} x_i &\sim \text{Gamma}(3, 0.2) \\ t_0 &\sim N(0, 100) \\ k &\sim U(0, 0.5) \\ \frac{1}{\sigma^2} &\sim \text{Gamma}(0.1, 0.1). \end{aligned}$$

Table 4 summarises the results of the simulation.

The results clearly show that ignoring the measurement error, even when the error is unbiased, will lead to biased parameter estimates. This is particularly noticeable in the parameters L_∞ and t_0. Also of interest is that ignoring the measurement error leads to larger estimates of σ_y, since this must incorporate uncertainty due to the measurement error as well as the uncertainty due to

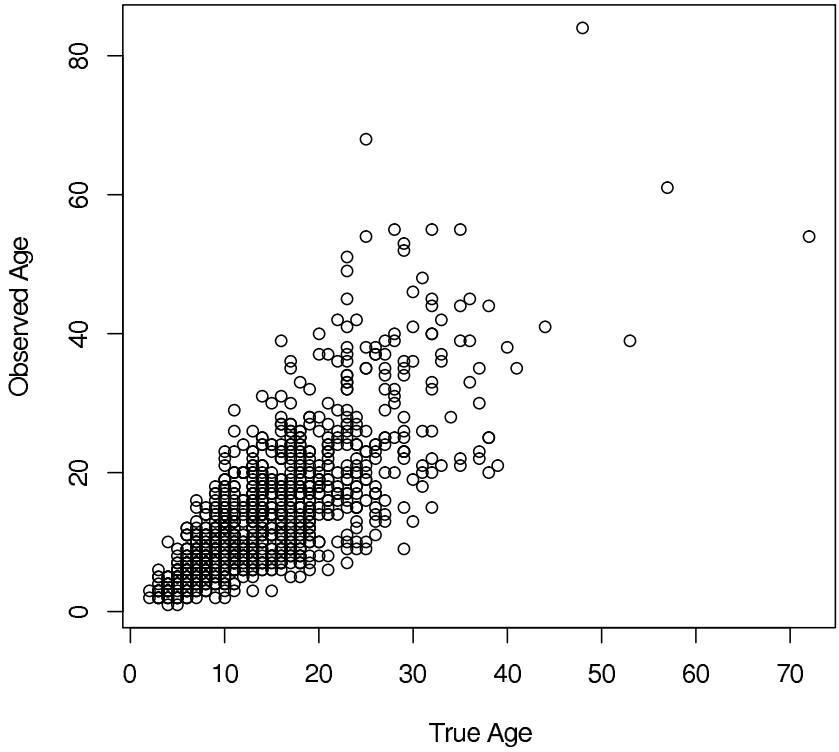

Figure 7: Simulated relationship between true age (x) and observed age (z).

Table 4: Results of Gibbs sampling for age/length simulation. All estimates are the means from each of 100 simulations. Mean standard deviations for each parameter are given in parentheses. Here *naïve* refers to the model in which measurement error is ignored.

Parameter	True	Naïve	ME
L_∞	130.00	125.45 (1.19)	129.83 (0.85)
k	0.15	0.14 (0.0099)	0.15 (0.0063)
t_0	-0.05	-3.59 (0.54)	-0.37 (0.23)
σ_y	7.00	13.87 (0.31)	7.15 (0.36)
λ	6.00	-	6.00 (0.23)

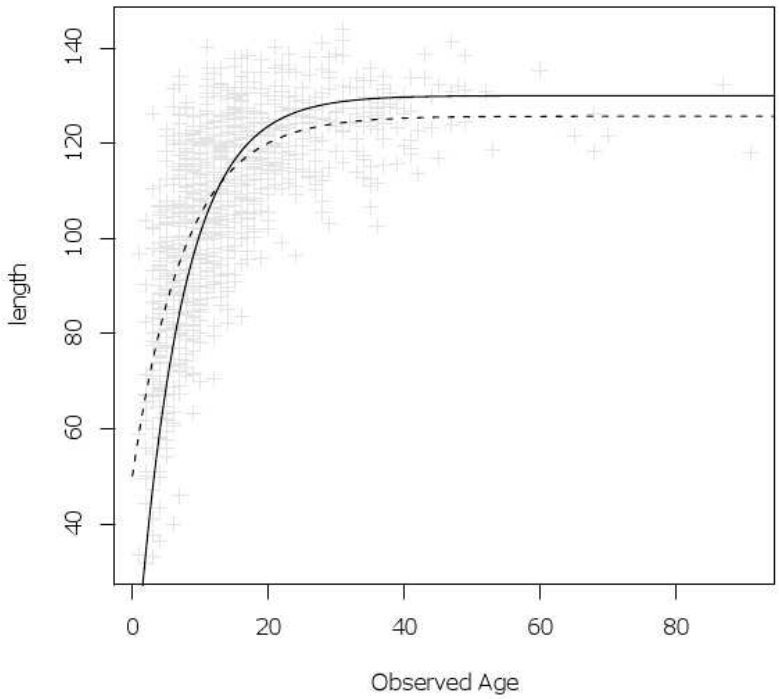

Figure 8: A simulated dataset. The unbroken line shows the true relationship between age and length, and the broken line shows the estimated relationship using observed age rather than true age.

the natural scatter around the line. So we can see that parameters are more accurate and precise when estimated using the measurement error model.

This example also illustrates one of the features of the Bayesian conditional independence model, in that the extension from linear models to non–linear is straightforward.

4 Discussion

The results from the examples suggest particular situations in which the ME model may be a useful approach when combining data. The situation in which the benefit from using the ME model is clearest is the case in which we have a small amount of accurate data, and access to a large amount of auxiliary data as in Section 3.1. In this case, we use the term auxiliary to describe data that is of itself of no real interest. To expand further, we are not interested in the relationship between the response and the auxiliary data, and we are not intending to make predictions of the response based on future observations of the auxiliary data. There are a number of situations in which this may occur within ecological data. A field survey may be designed specifically to relate a particular species' presence to micro–habitat variables. It may be possible to augment this survey data with earlier, more general fauna studies in which slightly different or less accurate micro–habitat variables were measured.

The species' distribution modelling example in Section 3.2 illustrated the need for strict adherence to the assumptions made in the ME model. In this case, conditional independence did not appear to be entirely appropriate, as reinforced by the likelihood ratio test. Additionally, there was considerable doubt over the adequacy of the validation data, both of which contributed to a model that yielded poor predictions. This highlights the fact that such models need to be used with caution, though of course this is not restricted to ecological applications. Phillips and Smith (1992) showed that even small errors in the assumptions in measurement error correction models can do more harm than good.

On the other hand, the problem of determining fish age to produce growth curves in Section 3.3 is an example in which ignoring the measurement error can have serious results. We are interested in accurate estimation of all the parameters involved in the model. This then will inform us of the true underlying relationship between age and fish length, rather than the surrogate (estimated age by otolith analysis). This is in contrast to many predictive models, including the species distribution example already discussed, where we were interested in prediction from the surrogate.

The specification of the measurement error problem under the Bayesian paradigm is simple and flexible, though not necessarily computationally straightforward, as evidenced by the slow convergence rate in even the simple simulation. This presents somewhat of a disincentive to widespread applications. It has been suggested that this is one reason that its application in measurement error problems, and particularly in epidemiology, has not been as widespread

as expected (Bashir and Duffy 1997). When applicable, though, the Bayesian Conditional Independence model has a number of advantages over standard regression approaches. It allows us to report on the features of interest, rather than on a surrogate, it provides a mechanism in which we can combine data of different qualities, and it allows us to correct biases in parameter estimates due to measurement error in the explanatory variables. There are situations in which the measurement error model is appropriate in natural resource problems, and methods, such as the one described in this paper, should be considered.

References

Adcock, R. J. (1877). Annote on the method of least squares. *Analyst 4*, 183–184.

Adcock, R. J. (1878). A problem in least squares. *Analyst 5*, 53–54.

Austin, M. P. and J. A. Meyers (1996). Current approaches to modelling the environmental niche of eucalypts—implications for management of forest biodiversity. *Forest Ecology and Management 85*, 95–106.

Bashir, S. A. and S. W. Duffy (1997). The correction of risk estimates for measurement error. *Annals of Epidemiology 7*, 154–164.

Beamish, R. J. and G. A. McFarlane (1983). The forgotten requirement for age validation in fisheries biology. *Transactions of the American Fisheries Society 112*(6), 735–743.

Berkson, J. (1950). Are there two regressions? *Journal of the American Statistical Association 45*, 164–180.

Campana, S. E. (2001). Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *J of Fish Biology 59*, 197–242.

Carroll, R. J., D. Ruppert, and L. A. Stefanski (1995). *Measurement Error in Nonlinear Models*. New York: Chapman and Hall.

Cheng, C.-L. and J. W. V. Ness (1999). *Statistical Regression with Measurement Error*. Kendall's Library of Statistics. Arnold.

Clayton, D. G. (1992). Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In J. H. Dwyer, M. Feinlieb, P. Lippert, and H. Hoffmeister (Eds.), *Statistical Models for Longitudinal Studies of Health*. New York: Oxford University Press.

Dellaportas, P. and D. A. Stephens (1995). Bayesian analysis of errors–in–variables regression models. *Biometrics 51*, 1085–1095.

Fernandes, R. and S. G. Leblanc (2005). Parametric (modified least squares) and non-parametric (Theil-Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment 95*(3), 303–316.

Fuller, W. A. (1987). *Measurement Error Models*. Wiley series in probability and mathematical statistics. Applied probability and statistics. New York: Wiley.

Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–511.

Gustafson, P. (2003). *Measurement Error and Misclassification in Statistics and Epidemology: Impacts and Bayesian Adjustments*. London: Chapman & Hall/CRC.

Hutchinson, M. F., H. A. Nix, D. J. Houlder, and P. McMahon, J (1998). *ANUCLIM Version 1.6 User Guide*. Canberra: Centre for Resource and Environmental Studies, The Australian National University.

Kuha, J. (1997). Estimation by data augmentation in regression models with continuous and discrete covariates measured with error. *Statistics in Medicine 16*, 189–20.

Mallick, B. K. and A. E. Gelfand (1996). Semiparametric errors-in-variables models. A Bayesian approach. *Journal of Statistical Planning and Inferernce 52*, 307–321.

Müller, P. and K. Roeder (1997). A Bayesian semiparametric model for case-control studies with errors in variables. *Biometrika 84*, 523–537.

Phillips, A. N. and G. D. Smith (1992). Bias in relative odds estimation owing to imprecise measurement of correlated exposures. *Statistics in Medicine 11*, 953–961.

Richardson, S. and W. R. Gilks (1993a). A Bayesian approach to measurement error problems in epidemiology using conditional independence models. *American Journal of Epidemiology 138*(6), 430–442.

Richardson, S. and W. R. Gilks (1993b). Conditional independence models for epidemiological studies with covariate measurement error. *Statistics in Medicine 12*, 1703–1722.

Spiegelhalter, D., A. Thomas, and N. G. Best (1996). *BUGS: Bayesian Inference Using Gibbs Sampling*. Cambridge: Medical Research Council Biostatistics Unit.

Williams, K., P. Norman, and K. Mengersen (2000). Predicting the natural occurence of blackbutt and gympie messmate in southeast queensland. *Australian Forestry 63*(3), 199–210.

Yuan, L. L. (2007). Effects of measurement error on inferences of environmental conditions. *Journal of the North American Benthological Society 26*(1), 152–163.

Zidek, J. V., H. Wong, N. D. Le, and R. Burnett (1996). Causality, measurement error and multicollinearity in epidemiology. *Environmetrics 7*(4), 441–451.